SSCI 583 Spatial Analysis

Project 3 Report

Group members: Junhong Duan, Lingduo Luo, Ruochen Liang, Shiqi Li, Yongchun Chen

1 Introduction

This project aimed to assess the accessibility of hospitals and clinics for elderly residents (age 65 and older) in San Francisco using the 2-Step Floating Catchment Area method (2SFCA).

As the population of elderly residents continues to grow, ensuring access to healthcare services is becoming increasingly important. Elderly residents may have unique healthcare needs and require more frequent medical attention, making accessibility to hospitals and clinics crucial. However, access to healthcare services can be influenced by various factors, such as distance to providers, transportation options, and provider availability. To better understand the accessibility of health care services for elderly residents in San Francisco, we conducted a spatial analysis incorporating demographic and geographic data. Our study provides a comprehensive picture of the healthcare accessibility landscape for elderly residents in San Francisco County. It can help identify areas where resources should be allocated to improve access to care. The study aimed to determine the extent to which elderly residents in San Francisco have access to health care services and identify any areas of the county where access may be limited.

To achieve this objective, we utilized a census data set that contained the numbers of elderly residents per block group and a data set that included the locations of hospitals and clinics in San Francisco. We calculated provider-to-population ratios and travel times from each block group to each hospital and clinic using ArcGIS Pro and the Network Analyst extension. We also created a catchment area for each healthcare provider, representing the geographic area where residents could access that provider in a reasonable amount of time. Finally, we combined these results to create an accessibility index, which reflected the overall level of accessibility to hospitals and clinics for elderly residents. The results of this study provide valuable insights into the spatial patterns of healthcare accessibility for elderly residents in San Francisco County and can inform policy decisions aimed at improving access to healthcare services for this vulnerable population.

2 Study Area

The study area of this project is San Francisco County, located in the northern part of California, United States. It is the most densely populated county in California and the second-most densely populated county in the United States, with a population of 873,965 in 2020.[1] San Francisco County covers an area of 121.4 square miles and has a diverse demographic composition, with about 15.9% of the population aged 65 years or older.[2] As is shown in Fig.1, the population of people aged 65 years or older in each census tract varies from 0 to 910. The census tracts with the most elderly population are in downtown San Francisco and the southwest of San Francisco County. The county is home to many major healthcare providers, including the University of California, San Francisco Medical Center, California Pacific Medical Center, and the San Francisco General Hospital.

San Francisco County is served by various modes of public transportation, including buses, light rail, cable cars, streetcars, and ferries. However, not all areas have equal access to these transit options, and some may face challenges in reaching health facilities due to physical barriers, traffic congestion, or lack of information.

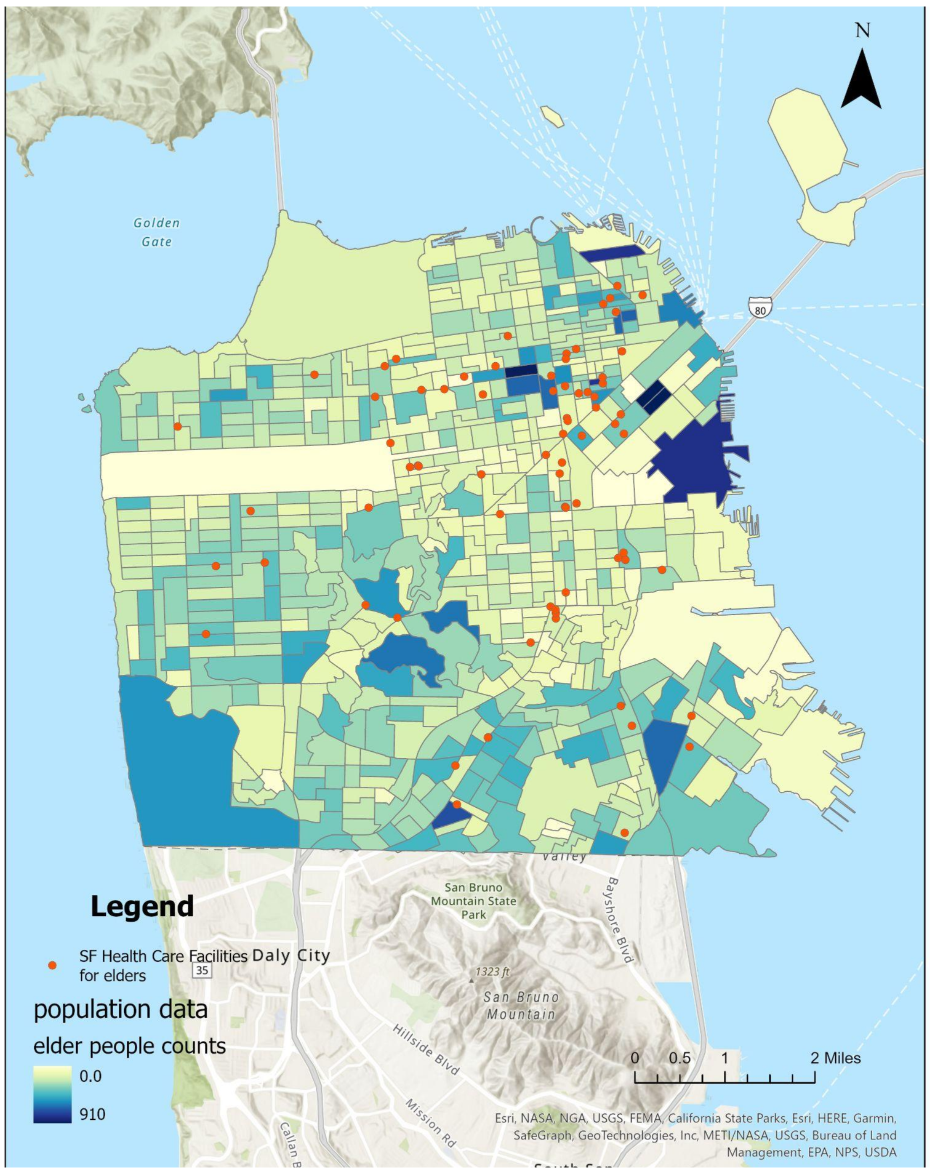

Fig.1 Study Area

3 Data and Data Preprocessing

3-1 Data Introduction

Table 1 shows that the dataset consists of three different types of spatial data related to San Francisco: census tract population, healthcare facilities, and a network. The census tract population data is sourced from the US Census and is represented as a vector polygon shapefile containing demographic information about the population in each census tract. The healthcare facilities data is sourced from DataSF and is described as a vector point shapefile containing information about the location and type of healthcare facilities in San Francisco. The network data is a vector line shapefile and is hosted as an ArcGIS Online Service, containing information about the road network in San Francisco.

Table.1 Data Table

Dataset	Source	File Type
SF census tract population	US Census	Vector Polygon Shapefile
SF healthcare facilities	DataSF	Vector Point Shapefile
SF network	ArcGIS Online Service	Vector Line Shapefile

3-2 Data Preparation

3-2-1 Filter out the elderly data over 65 years old

The SF census tract population dataset is a collection of polygon features representing US Census block groups. This dataset is supplemented with tabular data from the US Census that provides population numbers by sex and age for each block group. The purpose of this dataset is to provide demographic information for small geographic areas in San Francisco. The data can be used for research and analysis in various fields. We utilize this dataset to obtain the number of older adults in San Francisco. By the definition of Organisation for Economic Co-operation and Development[3], we define "elderly" as individuals who are 65 years old or above and then calculate the sum of the male and female populations separately to meet the needs of subsequent analysis better.

a) Three new columns were added, namely F_elder (total number of female elderly), M_elder (total number of male elderly), and M_F_elder (total number of elderly of both genders). Use the field calculator to add them as below.

- *M_elder = !M_65and66! + !M_67to69! + !M_70to74! + !M_75to79! + !M_80to84! + !M_85plus!*
- *F_elder = !F_65and66! + !F_67to69! + !F_70to74! + !F_75to79! + !F_80to84! + !F_85plus!*
- *M_F_elder = !M_elder! + !F_elder!*

3-2-2 Filter the suitable healthcare facilities for elderly people

a) The only data entry for a hospital catering to young adults was removed. (Female hospitals were not removed since they can cater to elderly women as well.)

b) Create a new "capacity" field in the SF census tract population dataset. Following Sarain's practice[4], we replaced the number of physicians with the capacity of the healthcare facility, which is the maximum number of patients that can be served daily. Assuming an average consultation time of one patient is one hour, and each clinic and health network has 40 working hours per week, their average daily capacity is eight patients, both for community clinics and healthcare networks. And for hospitals, the number of physicians in specialized healthcare facilities for elders is found directly on the hospital's website and multiplied by 16[5], the average number of patients each physician can take daily.

The data distribution after modification is shown below in Fig.2. This figure shows the distribution of female(left) and male(right) elders in San Francisco after data pre-processing. Also with the distribution of medical facilities all around the city. Image generated using ArcGIS Pro in March 2023.

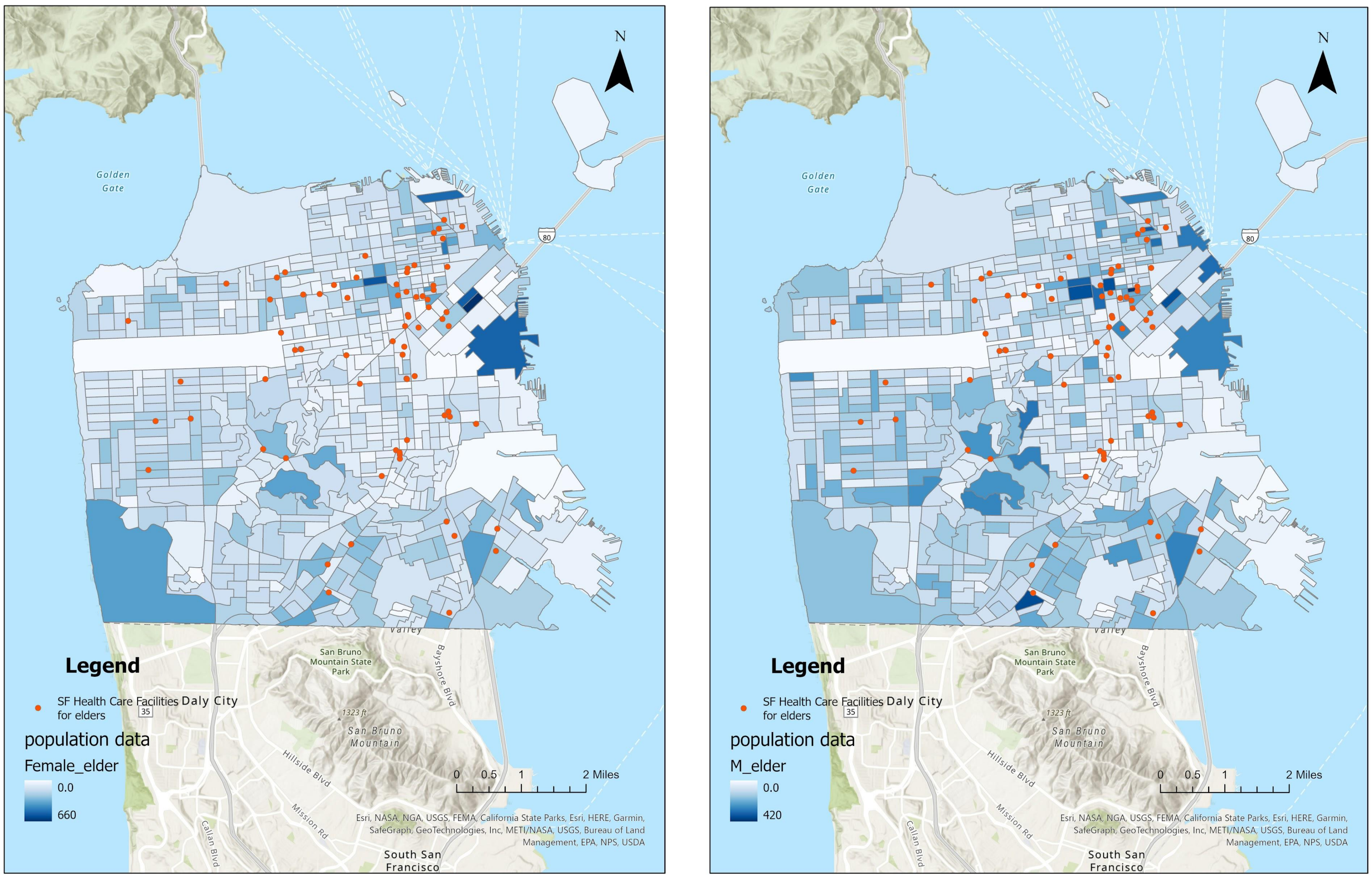

Fig.2 the distribution of female(left) and male(right) elders in San Francisco after data pre-processing

3-2-3 Transfer the SF census tract population dataset to points layer

By using the Feature To Point tool in ArcGIS Pro, we converted the census tract layer into a point layer with the centroid of each census tract polygon as the point location as in the figure below. The Inside parameter is not checked, so the output point will be located at the center of gravity of the polygon. In Figure 3, the color scheme is quintile, in which darker blue represents census blocks with more older people, and lighter blue represents census blocks with less older people. We could notice that older people are more likely to live in the northeastern part of San Francisco (Western Addition, Mission DIstrict, and Potrero District). Sunset District and Fort Winfield Scott have low density of older people.

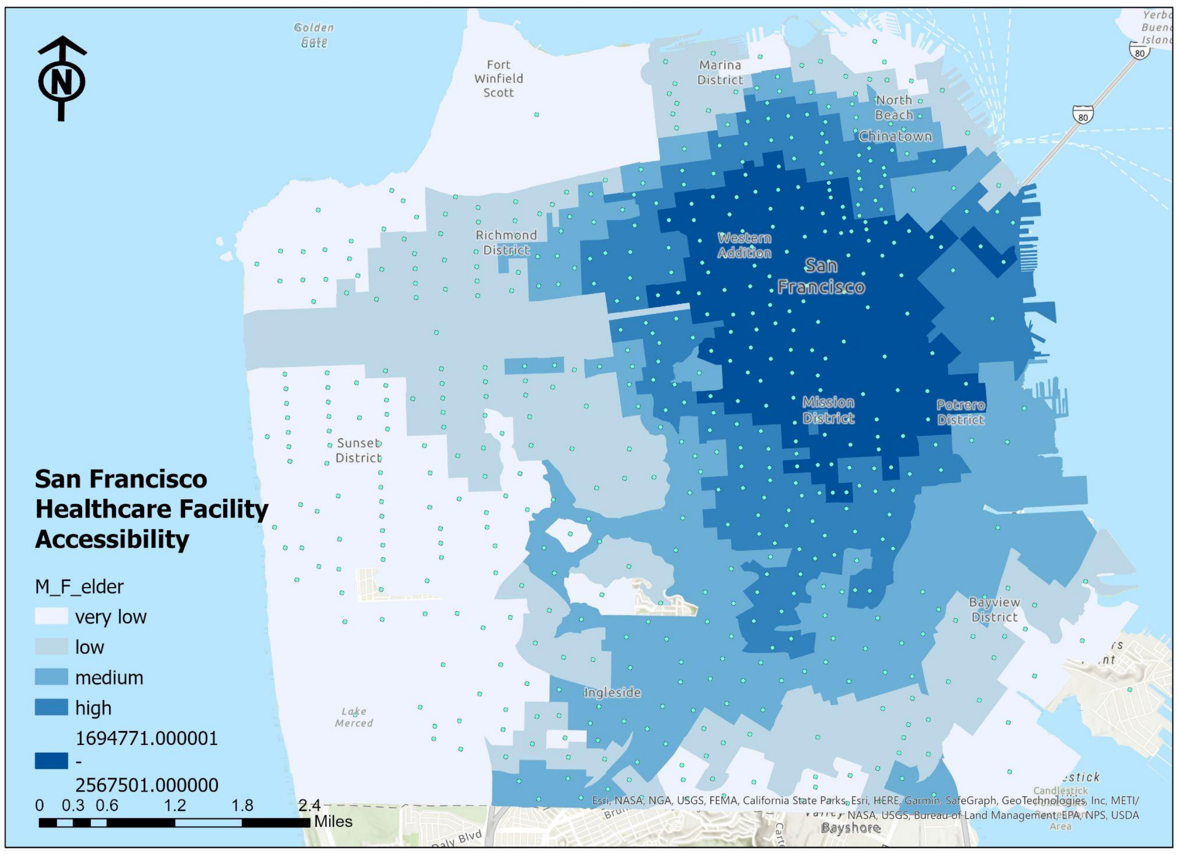

Fig.3 Census Tract Centroid Dataset with the elder population

These preprocessing steps were performed to ensure that the population and hospital data are relevant to the study of hospital accessibility for the elderly population in San Francisco.

4 Method

The 2-Step Floating Catchment Area (2SFCA) method is a popular technique used in spatial analysis to analyze accessibility to various amenities, including healthcare, transportation, and education[6]. As Fig.4 shows, the 2SFCA method works by dividing a study area into a set of catchments, with each catchment representing a geographic area within which residents have access to a specific facility or service. In the first step of the analysis, the method calculates a ratio of providers to population for each catchment. In the second step, the ratios are summed across all the catchments to determine the overall level of accessibility for the entire study area.

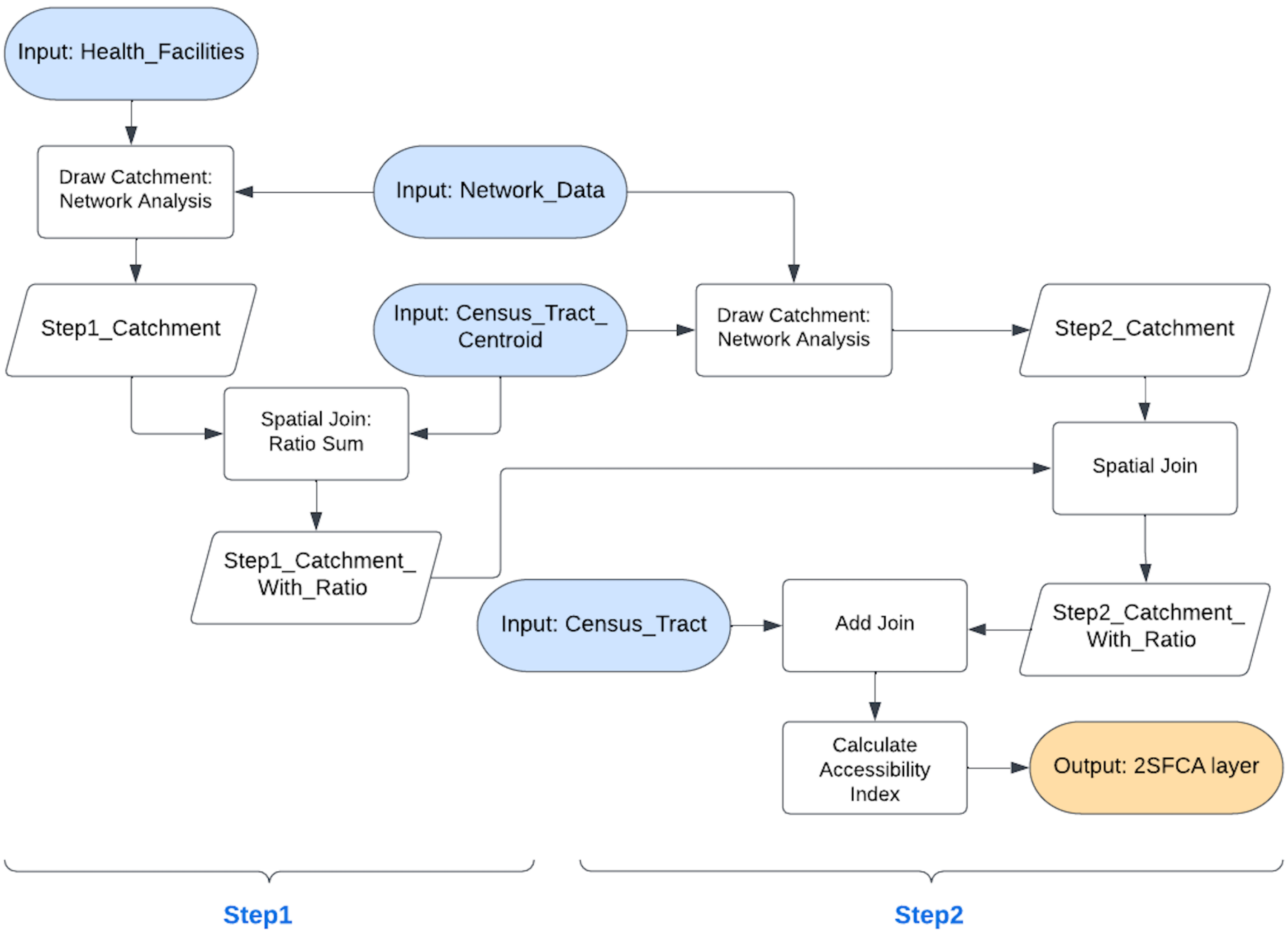

Fig.4 Workflow Diagram for Performing 2SFCA

4-1 2SFAC Step1

In this study, we defined the accessible area around each healthcare facility as the zone within a 10-minute driving distance. This catchment area was chosen based on the research of Luo and Qi[7] and aimed to reduce the 20-30 minute travel time zone resulting from the small study area. To create the catchment area, we utilized the Network Analysis tool in ArcGIS Pro, setting the network dataset as default from ArcGIS Online datasets and importing the elders' healthcare facility as the facility. The Mode was set to Driving time, the Direction was set to Towards Facilities, and the Cutoffs was set to 10. In the Output Geometry group, we changed Standard Precision to High Precision, consistent with the ArcGIS course's standard, Assess access to public transit[8], except for leaving the Overlap unchanged.

For step 1 of the 2SFCA and E2SFCA methods, the Assess ratio was calculated as the sum of all physician-to-population ratios. The physician-to-population ratio was computed by dividing the number of physicians by the population within the catchment area.

To compute the Capacity-to-Population ratio, we utilized the Spatial Join tool in ArcGIS Pro, setting the Join Operation to Join one to one because in this step, multiple join features are found that have the same spatial relationship with a single target feature, the study want the attributes from the multiple join features be aggregated using a field map merge rule[9], which is set to Sum to calculate the sum up elderly population in that catchment area. Therefore the Match Option is set to Contains. This calculation resulted in the sum of the elders' population within each catchment area. Following the Spatial Join, we created a new field named Ratio and used the Field Calculator in ArcGIS Pro to calculate the Capacity-to-Population ratio. The equation of step one is shown in the figure below.

Step 1 : Health centre (HC) to population (P) ratio

For each health centre location j, search for all the population or habitation locations (k) that are within the threshold travel distance (d_0) from a given location j and the computation of heath care to population ratio, R_j, within the catchment area:

$$R_j = \frac{HC_j}{\sum_{k \in \{d_{kj} \le d_0\}} P_k} \tag{1}$$

Fig.5 Formula of Step 1 to Calculate Provider to Population Ratio in E2SFCA[7]

4-2 2SFAC Step2

Step two of our analysis involves aggregating the provider ratios for each catchment area around a census tract centroid. To accomplish this, we first draw a catchment area around each centroid using the service area tool in the Network Analysis extension of ArcGIS. We set the drive time to 10 minutes, with the direction set as "away from facilities" and the result type set as "overlap". This creates six elements in each service area layer, with facilities represented as census tract centroids and polygons representing the corresponding catchment areas. Once these catchment areas are created, we use a spatial join to link them with the healthcare facilities dataset. This allows us to sum the provider ratios (R) for each healthcare facility within a given catchment area, resulting in a value named A(Accessibility Index of population, the formula is shown below) for each census tract. Aggregating the provider ratios in this way allows us to assess the overall accessibility to healthcare services for residents of each census tract, providing valuable insights for policymakers and urban planners seeking to improve healthcare equity in the community. Census tract maps will be applied by the quintile color scheme to show which polygons have higher or lower accessibility to healthcare services.

Step 2: Accessibility Index of population

$$A_i^F = \sum_{j \in \{d_{ij} \le d_0\}} R_j \tag{2}$$

P_k	the population at location k whose centroid falls within the catchment ($d_{kj} \le d_0$).
HC_j	the number of physicians at location j, and
d_{kj}	the travel distance between k and j.
A_i^F	the accessibility of population at a given location i to health care based on the two step floating catchment area method. The superscript F represents it's the calculation formula based on 2SFCA method.
R_j	health care to population ratio at location j whose centroid falls within the catchment ($d_{ij} \le d_0$).
d_{ij}	the travel distance between i and j.

Fig.6 Formula of Step 2 to Calculate Accessibility Index of Population in E2SFCA[7]

5 Result

In the final step, an "add join" function is employed to link column A to the census tract data, which is stored in a polygon shapefile format. To visualize the accessibility distribution throughout the study area, we utilize the symbology function in ArcGIS Pro to assign different colors based on the values in column A of the census tract polygon layer, resulting in the production of an accessibility map. Regions that exhibit darker shades on the map represent the geographic areas where the elderly population can conveniently access healthcare resources such as hospitals, clinics, and healthy networks. On the other hand, areas with lighter shades indicate that the elderly inhabitants of that region have limited access to such facilities.

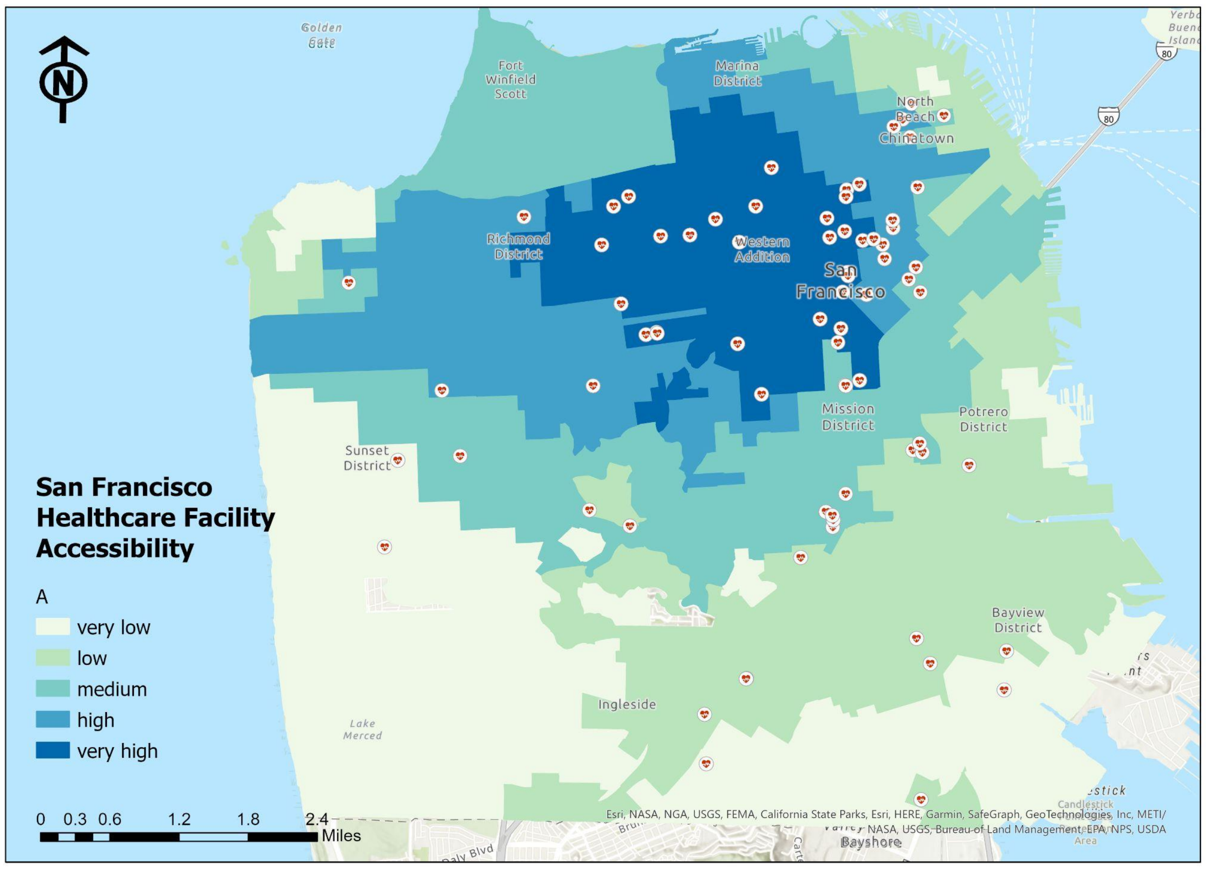

Fig.7 Final Result Map

As is shown in Fig.7, we could learn that census tracts in the north will have higher accessibility, especially in Western Addition and parts of Richmond District. However, the southern part of San Francisco will have very low accessibility because of the lack of the clinics and high-supply-capacity hospitals (Sunset District and Bay view District). Therefore, based on our analysis, it can be concluded that the accessibility of medical institutions for elderly residents in San Francisco is generally more convenient within a 10-minute driving time. Places with higher density of old people will be satisfied by the current healthcare supply capacity.

Consequently, to offer appropriate recommendations for medical institutions catering to the elderly population in San Francisco, we suggest that new medical institutions be established in the west and southeast, or alternatively, the capacity of existing medical institutions in this area should be increased.

6 Discussion

6-1 Meaning in the real world

This network analysis project, which conducts healthcare facilities network analysis with census tract data for people over 65 years old in San Francisco, California, can help identify areas with a high concentration of older adults. According to the San Francisco Department of Public Health, approximately 15% of the city's population is over the age of 65, and there would be nearly 30% of San Francisco residents will be 60 or older by 2030[9]. For example, the Outer Richmond and Outer Sunset neighborhoods have a higher proportion of older adults than the citywide average. Analyzing healthcare facility data in these areas can help identify where additional resources may be needed to meet the healthcare needs of older adults.

Healthcare facility network analysis can also assess access to care and address health disparities for older adults in San Francisco. The 2019 San Francisco Community Health Needs Assessment found that older adults in the city face several barriers to accessing healthcare, including transportation, cost, and language barriers. Analyzing healthcare facility data alongside census tract data can help identify areas where older adults may be facing these barriers to care.

Additionally, San Francisco has significant health disparities across neighborhoods, with life expectancy varying by up to 14 years, depending on the community. Analyzing healthcare facility data alongside census tract data can help identify areas where older adults may face health disparities and where additional resources may be needed to address these disparities.

Generally, healthcare facilities network analysis with census tract data can help identify areas where additional resources are needed to meet the healthcare needs of older adults, assess access to care, and address health disparities.

MAUP means the Modifiable Areal Unit Problem, which refers to the phenomenon where statistical analysis results can be influenced by choice of geographic units used to aggregate data. The problem arises when the same data is analyzed at different spatial scales, resulting in different conclusions about relationships between variables. This is because the choice of geographic units can affect the magnitude and direction of spatial patterns in the data. In this study, the catchment area for both census tract centroid and healthcare facilities is defined as a 10-minute drive zone because San Francisco is a densely populated city with narrow streets and many one-way streets, making travel time by car a more reliable measure than distance. A 10-minute driving zone provides a more accurate estimate of travel time than other distance-based measures, such as a 1-mile radius, which may not accurately reflect driving conditions in the city.

6-2 Limitations

San Francisco is a densely populated city with narrow streets, heavy traffic, and limited parking, making it challenging to get around by car. Walking distance or public transportation time is a more accurate measure of geographic proximity in San Francisco than driving distance, particularly in areas with high pedestrian traffic and limited parking. Moreover, many residents of San Francisco do not own a car and rely on public transportation, walking, or biking to get around the city. By using walking distance, likely, people can better understand the accessibility of healthcare facilities for people who do not have access to a car. So, choosing the mode of transportation and type of catchment areas is the project limitation.

7 Reference

[1] “U.S. Census Bureau Quickfacts: San Francisco County, California.” Accessed March 11, 2023. https://www.census.gov/quickfacts/sanfranciscocountycalifornia.

[2] “San Francisco.” Wikipedia. Wikimedia Foundation, January 28, 2023. https://en.wikipedia.org/wiki/San_Francisco.

[3] https://data.oecd.org/pop/elderly-population.htm

[4] Sarain, L. (2019). Providing A New Low-Cost Primary Care Facility for Under-Served Communities: A Site Suitability Analysis for Service Planning Area 6 in Los Angeles County, California [MA thesis]. University of Southern California. https://spatial.usc.edu/wp-content/uploads/formidable/12/Yue-Li.pdf

[5] “How Many Patients Does a Doctor See Daily? | Elation Health EHR.” Www.elationhealth.com, 25 July 2022, www.elationhealth.com/starting-a-new-practice-blog/how-many-patients-does-a-doctor-have-a-day/. Accessed 12 Mar. 2023.

[6] Wang F. From 2SFCA to i2SFCA: integration, derivation and validation[J]. International Journal of Geographical Information Science, 2021, 35(3): 628-638.

[7] Luo, Wei, and Yi Qi. “An Enhanced Two-Step Floating Catchment Area (E2SFCA) Method for Measuring Spatial Accessibility to Primary Care Physicians.” Health & Place 15, no. 4 (2009): 1100–1107. https://doi.org/10.1016/j.healthplace.2009.06.002.

[8] “Assess Access to Public Transit.” Learn ArcGIS. Accessed March 5, 2023. https://learn.arcgis.com/en/projects/assess-access-to-public-transit/.

[9] “Spatial Join (Analysis).” Spatial Join (Analysis)-ArcGIS Pro | Documentation. Accessed March 11, 2023. https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/spatial-join.htm.

8 Appendix

Table. 2 Individual Contribution

Name	Practice	Report
Junhong Duan	2SFCA Method Performing, Network Analysis, Data Searching, Result Generation	Method and Result
Lingduo Luo	Data Preparation, Literature Review, Map Generation, Workflow Presentation	Data and Data Preprocessing
Ruochen Liang	Data Searching, Literature Review, Network Analysis, Result Check	Method and Discussion
Shiqi Li	Literature Review, Network Analysis, Result Check, Data Preparation	Method and Study Area
Yongchun Chen	Workflow Presentation, Literature Review, Data Preparation, Map Generation	Introduction, Method and Result